

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup>:</b> <b>C12N 15/66, C12Q 1/68</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 97/04111</b> <b>(43) International Publication Date:</b> 6 February 1997 (06.02.97)
<b>(21) International Application Number:</b> PCT/US96/11721 <b>(22) International Filing Date:</b> 16 July 1996 (16.07.96)  <b>(30) Priority Data:</b> 60/001,384      21 July 1995 (21.07.95)      US  <b>(71) Applicant:</b> THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Office of Technology Transfer, Suite 325, 6011 Executive Boulevard, Rockville, MD 20852-3804 (US).  <b>(72) Inventors:</b> FERRIN, Lance, Joseph; 8800 Alliston Hollow Way, Gaithersburg, MD 20879 (US). CAMERINI-OTERO, Rafael, Daniel; 3619 Littledale Road, Kensington, MD 20795 (US).  <b>(74) Agent:</b> ALTMAN, Daniel, E.; 16th floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>

**(54) Title:** RecA-ASSISTED CLONING OF DNA**(57) Abstract**

DNA is cloned and labeled in a sequence-specific manner. The DNA is digested with one or more restriction enzymes which produce 3' recessed ends. A desired fragment is protected from elongation by DNA polymerase by addition of *E. coli* RecA protein and oligonucleotides about 30 to 60 bases in length complementary to the 3' recessed ends of the digested fragment. RecA and DNA polymerase are then inactivated, leaving only the desired fragment with 3' recessed ends which is then ligated into a vector containing complementary 3' recessed ends.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland			US	United States of America
FR	France				

## RecA-ASSISTED CLONING OF DNA

### Field of the Invention

The present invention relates to the sequence-specific cloning and labeling of DNA using the RecA protein. More specifically, the invention relates to the ability of RecA to selectively pair oligonucleotides to their homologous sequences at the 3' recessed ends of digested duplex DNA fragments and to protect these 3' ends from enzymatic conversion to blunt ends, thus facilitating cloning of a desired DNA fragment.

### Background of the Invention

The isolation and cloning of genomic DNA fragments is of paramount importance to the biomedical sciences. In this regard, several methods are available to amplify DNA and to isolate selected fragments in pure form. The most widely used amplification method is the polymerase chain reaction (PCR). In this method, oligonucleotide primers flanking a desired DNA sequence are used to amplify the sequence by repeated rounds of denaturation, annealing and extension steps. However, a major limitation of PCR is the small fragment size which may be reliably amplified, although recent improvements have allowed amplification of up to 22 kilobases (kb) (Cheng et al., *Proc. Natl. Acad. Sci. USA*, 91:5695, 1994; Foord et al., *PCR Methods and Applications*, 3:S149, 1994).

Other widely used methods of cloning genomic DNA fragments involve the construction and screening of DNA libraries, most commonly  $\lambda$  phage and cosmid vectors. Other vectors are now gaining widespread use for cloning large (> 100 kb) segments of DNA including yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs) and P1 phage derived artificial chromosomes (PACs). Such libraries, however, are difficult to construct and screen.

*E. coli* RecA protein has been used to screen libraries and to enrich for a selected DNA fragment (Rigas et al., *Proc. Natl. Acad. Sci. USA*, 83:9591, 1986; Honigberg et al., *Proc. Natl. Acad. Sci. USA*, 83:9586, 1986; Taidi-Laskowski et al., *Nucl. Acids. Res.*, 16:8157, 1988; Sena et al., *Nature Genet.*, 3:365, 1993; Jayasena et al., *J. Mol. Biol.*, 230:1015, 1993). These methods are based on the ability of RecA to specifically target single-stranded DNA to complementary target duplex DNA to create a three-stranded complex (Camerini-Otero et al., *Cell*, 73:217, 1993), or to pair two complementary single strands to the target duplex DNA to create a four-stranded complex. These strategies have not been applied to practical problems in molecular biology.

RecA-Assisted Restriction Endonuclease (RARE) Cleavage is a general and efficient method of targeting restriction enzyme cleavage to unique predetermined sites and is described in U.S. Patent Application Serial No. 08/089,910, the entire contents of which are hereby incorporated by reference, and by Ferrin et al. (*Nature Genet.*, 6:379, 1994). This method is based on the ability of RecA to pair oligonucleotides to homologous sequences in duplex DNA to form three-stranded complexes. These complexes protected the selected sites from enzymatic manipulation, and, after removal of the complexes, restriction enzyme cleavage was limited to the selected unmethylated sites. This method has been used to map and manipulate large segments of DNA (Ferrin, in *Genetic Engineering: Principles and Methods*, J. Setlow, Ed., Plenum Press, New York, 17:21-30, 1995; Barton et al., *Genes and Dev.*, 8:2453, 1994; Heineman et al., *J. Virol.*, 68:3317, 1994; Gourdon et al., *Nucl. Acids. Res.*, 22:4139, 1994).

material includes instructions for using the reagents to clone DNA, such as by the method described above. The reagents can also include one or more restriction enzymes capable of generating 3' recessed ends, DNA polymerase and a vector having 3' cohesive ends. The 3' cohesive ends are preferably complementary to the 3' recessed ends generated by the restriction enzymes. In one embodiment, these restriction enzymes are *EcoRI* and *BamHI*. The DNA polymerase can be the Klenow fragment of *E. coli* DNA polymerase I, and the vector can be a plasmid.

#### Brief Description of the Drawings

Figure 1 is a schematic diagram of the strategy used for sequence-specific RecA-mediated amplification of DNA. The RecA-oligonucleotide complexes are indicated by the adjacent circles having a line passing through their centers.

Figure 2 is a schematic diagram of  $\lambda$  DNA digested with *HindIII*. The 2.3 kb fragment labeled by RecA-Assisted cloning is indicated by the arrow. The left (L) and right (R) 3' ends to which the L and R oligonucleotides are complementary are shown.

Figure 3 is a schematic diagram of the construct resulting from ligation of the 1.4 kb human *int-2* genomic DNA fragment to a vector and to a biotinylated DNA duplex bound to a streptavidin-coated bead. *EcoRI* and *BamHI* restriction sites and 4 base pair 5' overhangs are shown. S, streptavidin; B, biotin; and P, 5' phosphate.

#### Detailed Description of the Preferred Embodiments

The present invention relates to a method of sequence-specific genomic DNA cloning. This method is based on the ability of *E. coli* RecA protein to selectively pair oligonucleotides to their complementary sequences at the ends of duplex DNA. Genomic DNA is digested with one or two restriction enzymes which produce 3' recessed ends (5' overhangs). After addition of RecA protein and a pair of oligonucleotides, each complementary to one of the ends of a genomic DNA fragment of interest, the resulting three-stranded complexes become resistant to elongation by DNA polymerase and thus retain their 3' recessed ends after addition of the enzyme, while the unprotected genomic DNA fragments are filled in by the polymerase, thus becoming blunt-ended. Because most restriction endonucleases produce fragments having 3' recessed ends, these fragments were targeted for amplification using the method of the present invention. By using ligation conditions and vectors which greatly favor ligation to 3' recessed ends, protected fragments are selectively cloned. The vector into which the genomic DNA fragment is to be inserted is digested with the same restriction enzyme(s) as was the genomic DNA, or with a restriction enzyme(s) that produce the same 5' overhangs, resulting in complementary 3' recessed ends for insertion of the genomic DNA fragment.

The oligonucleotides used in the present method are complementary to a portion of the ends of a desired DNA fragment, including the 5' overhangs themselves. For example, digestion with *EcoRI* produces the 5' overhang TTAA. Thus, the oligonucleotides are complementary to this sequence plus additional sequence of the genomic DNA fragment adjacent to the overhang which is complementary to the remainder of the oligonucleotide. It will be appreciated that an oligonucleotide complementary to one DNA strand is identical in sequence to the other DNA strand. Of course, the oligonucleotides can be complementary to either strand at the end of the DNA duplex. Further, the oligonucleotides can also include additional nonhomologous sequences at their 5' or 3' ends; although these "nonhomologous tails" do not ordinarily increase the efficiency of protection. In addition, it is contemplated that only a single end of a desired DNA fragment can be protected with an oligonucleotide complementary thereto. In fact, a 200-fold enrichment of a particular

adenoviral vectors and the like. Particularly preferred vectors for the propagation of large DNA fragments include YACs, BACs and PACs. The vector is then used to transform cells which are expanded, resulting in amplification of the selected DNA fragment. Alternatively, the fragment can be amplified using PCR.

RecA-Assisted cloning has sufficient specificity to allow cloning directly from genomic DNA and is a much easier alternative than construction and screening of DNA libraries. The technique is preferable to PCR in the cloning of large (greater than about 5 kb) or highly repetitive fragments of DNA, especially if absolute fidelity is required due to the lower error rate of RecA-Assisted Cloning versus PCR.

RecA-Assisted Cloning can be used to identify specific mutations in a gene which give rise to genetic abnormalities and thus is useful in screening patients for genetic abnormalities or mutations which will predispose patients to certain conditions. Such mutations include point mutations, insertions and deletions. One particular use in this regard is in fetal screening. Fetal cells can be obtained by amniocentesis and analyzed for genetic defects including Tay-Sachs, sickle cell anemia,  $\beta$ -thalassemias, and any other desired genetic disease. Specific oligonucleotides are designed which will hybridize to the 3' ends of the fragment containing the DNA sequence of interest.

Many modifications of RecA-Assisted Cloning are contemplated. For example, the RecA/KF reactions worked well on DNA embedded in agarose which will be useful for molecules that would tend to shear in solution. For applications in which increased specificity is desired, RecA-Assisted Cloning can be used after RARE cleavage, or with type IIs restriction enzymes that create varied and asymmetric staggered ends unrelated to their recognition sites (Berger, *Anal. Biochem.*, 222:1, 1994). Increases in specificity would also be useful for labeling specific genomic DNA fragments using RecA-Assisted Cloning and is a viable alternative to detection methods such as Southern blotting. In addition, if conditions can be found that allow labeling of very short duplexes, the method can be a useful adjunct to sequencing by oligonucleotide array methods (Drmanac et al., *Science* 260:1649, 1993).

Sequence-specific labeling of a  $\lambda$  DNA fragment using RecA-Assisted Cloning was performed as described below.

#### Example 1

##### Sequence-specific labeling of a 2.3 kb $\lambda$ DNA fragment

*E. coli* RecA protein was prepared as described (Ferrin et al., *Science*, 254:1494, 1991) using an overproducing strain provided by Barbara McGrath of the Brookhaven National Laboratory, or purchased from Boehringer Mannheim (Indianapolis, IN). The sequence of the L oligonucleotide was 5'-gattatAGCTTTTCTAATTTAACCTTTGTCAAGGTTACCA-3' (SEQ ID NO:1), and the R oligonucleotide was 5'-gattatAGCTTTGTGTGCCACCCACTACGACCTGCATAA-3' (SEQ ID NO:2). The lower case letters indicate sequences of nonhomologous tails, and the capital letters indicate the sequences of portions homologous to the ends of the  $\lambda$  fragment. Oligonucleotides over 30 bases in length were purified on acrylamide gels and concentrations were measured as described (Ferrin et al., *supra*).

The RecA protection reaction volume was 100  $\mu$ l and contained 25 mM Tris-acetate, pH 7.85, 4 mM magnesium acetate, 0.4 mM dithiothreitol, 0.5 mM spermidine, 1.1 mM ADP, 0.3 mM ATP- $\gamma$ -S (Fluka), 13  $\mu$ g of RecA protein, 0.32  $\mu$ g L or R oligonucleotide (or 0.16  $\mu$ g each of L and R), 2.5  $\mu$ g of *HindIII*-digested  $\lambda$  DNA (New England Biolabs, Beverly, MA) and 40  $\mu$ g bovine serum albumin (BSA; Sigma, St. Louis, MO), 38  $\mu$ M each of dATP, dCTP, dGTP and TTP, and 12.5 units of KF (United States Biochemical, Cleveland, OH). After a 10 minute incubation at 37°C, KF

## Example 2

RecA-assisted cloning of genomic DNA

The human *int-2* proto-oncogene has been mapped and sequenced (Casey et al., *Mol. Cell. Biol.*, 6:502, 1986; Brookes et al., *Oncogene*, 4:429, 1989). In this gene, one *EcoRI* site lies just upstream of exon 2 and, in about half of the alleles, a *BamHI* site is 6.9 kb downstream of the *EcoRI* site. Human genomic DNA was isolated from multiple placentas (Sigma), digested with *EcoRI* and *BamHI*, extracted with phenol/chloroform and ethanol precipitated as described in Example 1. Yields after the ethanol precipitation were typically about 60%. Digested DNA was size fractionated on a 0.8% SeaPlaque GTG (FMC BioProducts) agarose gel in TAE buffer (Sambrook et al., *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory Press, Plainview, NY, Second Edition, 1989). Multiple wells were loaded with 150-200  $\mu$ g DNA per well. The gels were run until the 1.4 kb fragment had migrated 4 to 6 cm. The marker lane was then removed, stained with ethidium bromide and used as a guide to excise 0.5 cm above and below the position of the 1.4 kb fragment. Size fractionation resulted in a modest sequence enrichment (about ten fold); a ten-fold decrease in the amount of reagents required; and eliminated small fragments that would preferentially be represented in the final clones.

DNA was extracted from the excised gel using GELase<sup>®</sup> (Epicentre) according to the manufacturer's directions; complete digestion was required for good yields. The yield of the complete protocol was 2 to 4%. Comparable yields were obtained with a silica gel extraction kit (Qiagen) or by electroelution (Pun et al., *Prep. Biochem.*, 20:123, 1990). The size of the extracted DNA from the heavily overloaded gels was checked on analytical gels. Depending on the amount available, DNA was quantified by absorbance, fluorescence (Labarca et al., *Anal. Biochem.*, 102:344, 1980), or spotting in an ethidium bromide solution (Sambrook et al., 1989).

The size-fractionated doubly digested human placental DNA was used as the starting DNA for the RecA protein/KF protection reaction. The conditions for this reaction were the same as described in Example 1 with the exception that the total volume was 1440  $\mu$ l and contained 3.2  $\mu$ g of each oligonucleotide, 360  $\mu$ g RecA protein, 2.6  $\mu$ g fractionated DNA, 570  $\mu$ g BSA and 450 units of KF. One nucleotide was identical to the *int-2* genomic sequence from 2290-2347: 5'-GGTCCGAGTGC GCGGAATTCGTCTCACTAAGACACTCCGGTTCTCTCCAAAGCCAGGC-3'(SEQ ID NO:5), and the other was complementary to 3621-3677: 5'-TGGTCCTAGCTTGGATCCCATGTACCCTTGGCAAAGCATTCTACTGCCCACATCCCC-3'(SEQ ID NO:6). *EcoRI* and *BamHI* cleave 3' of residues 2304 and 3660, respectively (Casey et al., 1986; Brookes et al., 1989).

The protected fragments were ligated both to the pBS SK<sup>+</sup> vector (Stratagene, La Jolla, CA) and to DNA bound to streptavidin beads using T4 DNA ligase. This step reduces the number of clones containing only vector DNA. When plasmid or  $\lambda$  vectors were simply ligated to DNA from the RecA protein/KF reaction, the vast majority of clones did not contain an insert. Efforts were made to reduce this background by decreasing the vector concentration, but this also lowered the efficiency of the cloning procedure. Due to the low amount of the selected fragment in genomic DNA, a large concentration of vector facilitated the intermolecular vector-fragment ligation.

Ligation of the 1.4 kb fragment to the vector and to the biotinylated DNA duplex bound to the streptavidin-coated beads is schematically shown in Figure 3. Magnetic streptavidin beads (2 mg; Dynabeads M-280, Dynal) were

## Example 3

Multiple Cloning Trials

Multiple cloning trials were performed using both yeast and human DNA. With human DNA, the typical enrichment was 1,000 to 2,000-fold, and one *int-2* clone was present for every 2,000 to 4,000 colonies. At least one *int-2* clone was obtained for every 70  $\mu$ g of starting genomic DNA. When the pooled DNA after one round of RecA-Assisted Cloning was subjected to the procedure a second time, 24% of the colonies contained the *int-2* fragment. The procedure was essentially identical to that described in Example 2, except that the RecA/KF reaction step was scaled down by a factor of 14. pBC SK<sup>+</sup> was used as the vector and clones were selected on chloramphenicol plates to eliminate any background from the previous plasmid vector which contained an ampicillin resistance gene. This demonstrated an additional 500-fold enrichment and showed that incorrect clones arose mainly through a stochastic process, and not through a biased selection based on partial homology to the *int-2* sequence. A 1.2 kb *EcoRI-BamHI* yeast genomic DNA fragment containing the proximal portion of the RAD51 gene (Shinohara et al., *Cell* 69:457, 1992) was also cloned. The oligonucleotides used to clone this fragment had sequences complementary to positions 1-48: 5'-TGAAAATATTGAACAGTGAATAAAGCATAAAAAAAATGTCGGATCCATAGCGCTAT-3'(SEQ ID NO:9), and 1164-1204: 5'-GGACTTACCTGTCCTGTCCTGAATTCACCGAAAAGCTCAGTAATAGAACCAGTTTCCACACC-3'(SEQ ID NO:10). The yeast genomic DNA was isolated as described by Ausubel (*ibid.* (p.13)). Conditions were identical to those described in Example 2, except that the RecA/KF reaction was scaled down by a factor of 14, pBC SK<sup>+</sup> was used as the vector and clones were selected on chloramphenicol plates.

Plasmid DNA from 10 *int-2* clones and 10 RAD51 clones were analyzed by restriction enzyme mapping. No rearrangements were detected. The sequences of the two vector-insert junctions and about 400 bases of insert DNA of each of the 20 clones were determined. Plasmid DNA was prepared using a kit (Qiagen). Digestions were performed with *EcoRI* and *BamHI* and analyzed by electrophoresis on an agarose gel. Sequencing was performed on an Applied Biosystems model 373 sequencer using their PRISM DyeDeoxy Terminator Cycle Sequencing kit and the M13-20 and reverse primers. No clear deviations from the published sequences were detected, but as ambiguities in the sequences occurred at a rate of about 1%, this could only be used to set an upper limit to the error rate of RecA-Assisted Cloning. One might expect the error rate to be closer to the *in vivo* error rate in *E. coli* of  $10^{-10}$  mutations/bp/chromosome duplication (Schaaper et al., *J. Biol. Chem.*, 268:23762, 1993; Drake, *Proc. Natl. Acad. Sci. USA*, 88:7160, 1991), rather than the PCR error rate of about  $10^{-5}$  (Barnes, *Proc. Natl. Acad. Sci. USA*, 91:2216, 1994).

Although the invention has been described with reference to particular preferred embodiments, the scope of the invention is defined by the appended claims and should be construed to include reasonable equivalents.

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GATTATAGCT TTGTGTGCCA CCCACTACGA CCTGCATAA

39

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCCTACGAT CGATGCCTTG ACAT

24

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATGCTAGCTA CGGAAGTGA GGAG

24

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 58 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO



- (A) LENGTH: 58 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGAAAATATT GAACAGTGAA TAAAGCATAA AAAAAAATG TCGGATCCAT AGCGCTAT

58

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 57 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGACTTACCT GTCCTGAATT CACCGAAAAG CTCAGTAATA GAACCAAGTTT CCACACC

57

(b) digesting said genomic DNA with one or more restriction enzymes which generate 3' recessed ends to produce genomic DNA fragments having 3' recessed ends;

(c) reacting said genomic DNA fragments with RecA protein and two oligonucleotides, said oligonucleotides being complementary to the ends of the fragment containing the mutation;

5 (c) reacting the genomic DNA fragments resulting from step (b) with a DNA polymerase, whereby all genomic DNA fragments except the fragment containing the mutation become blunt-ended;

(d) dissociating said oligonucleotides from the ends of the fragment containing the mutation;

(e) ligating said DNA fragments to a vector having 3' recessed ends complementary to those produced by the restriction enzyme(s), whereby only the fragment containing the mutation is incorporated into said vector;

10 (f) amplifying the fragment containing the mutation; and

(g) determining whether the mutation is present in the amplified fragments.

16. The method of Claim 15, wherein step (f) comprises growth of said vector in a suitable microorganism.

17. The method of Claim 15, wherein step (f) comprises PCR.

18. The method of Claim 15, wherein said oligonucleotides are between 30 and 60 bases in length.

15 19. The method of Claim 15, wherein step (g) comprises sequencing said fragment.

20. The method of Claim 15, wherein step (d) comprises treating with sodium dodecyl sulfate or phenol/chloroform.

21. The method of Claim 15, wherein said mammal is a human.

22. The method of Claim 15, wherein said DNA polymerase is the exonuclease-deficient Klenow fragment  
20 of *E. coli* DNA polymerase I.

23. The method of Claim 15, wherein said vector is a plasmid.

24. The method of Claim 15, wherein said vector is a yeast artificial chromosome, bacterial artificial chromosome or P1 phage derived artificial chromosome.

25. The method of Claim 15, further comprising the step of size fractionating said DNA fragments of step  
25 (a) to enrich for the fragment containing the mutation.

26. The method of Claim 25, further comprising prior to the ligating step, ligating the enriched DNA fragments to a biotinylated duplex containing complementary cohesive ends, wherein said biotinylated duplex is bound to streptavidin-coated beads.

27. An article of manufacture comprising packaging material and at one or more reagents for cloning of  
30 DNA, wherein the reagents for cloning of DNA comprise recA, and wherein the packaging material comprises instructions for using the reagents to clone DNA.

28. An article of manufacture comprising packaging material and at one or more reagents for cloning of DNA, wherein the reagents for cloning of DNA comprise recA protein, and wherein the packaging material comprises instructions for using the reagents to clone DNA according to the method of Claim 1.

35 29. An article of manufacture according to Claim 27, wherein the instructions for using the reagents comprise instructions to identify one or more mutations in genomic DNA.

1 of 3

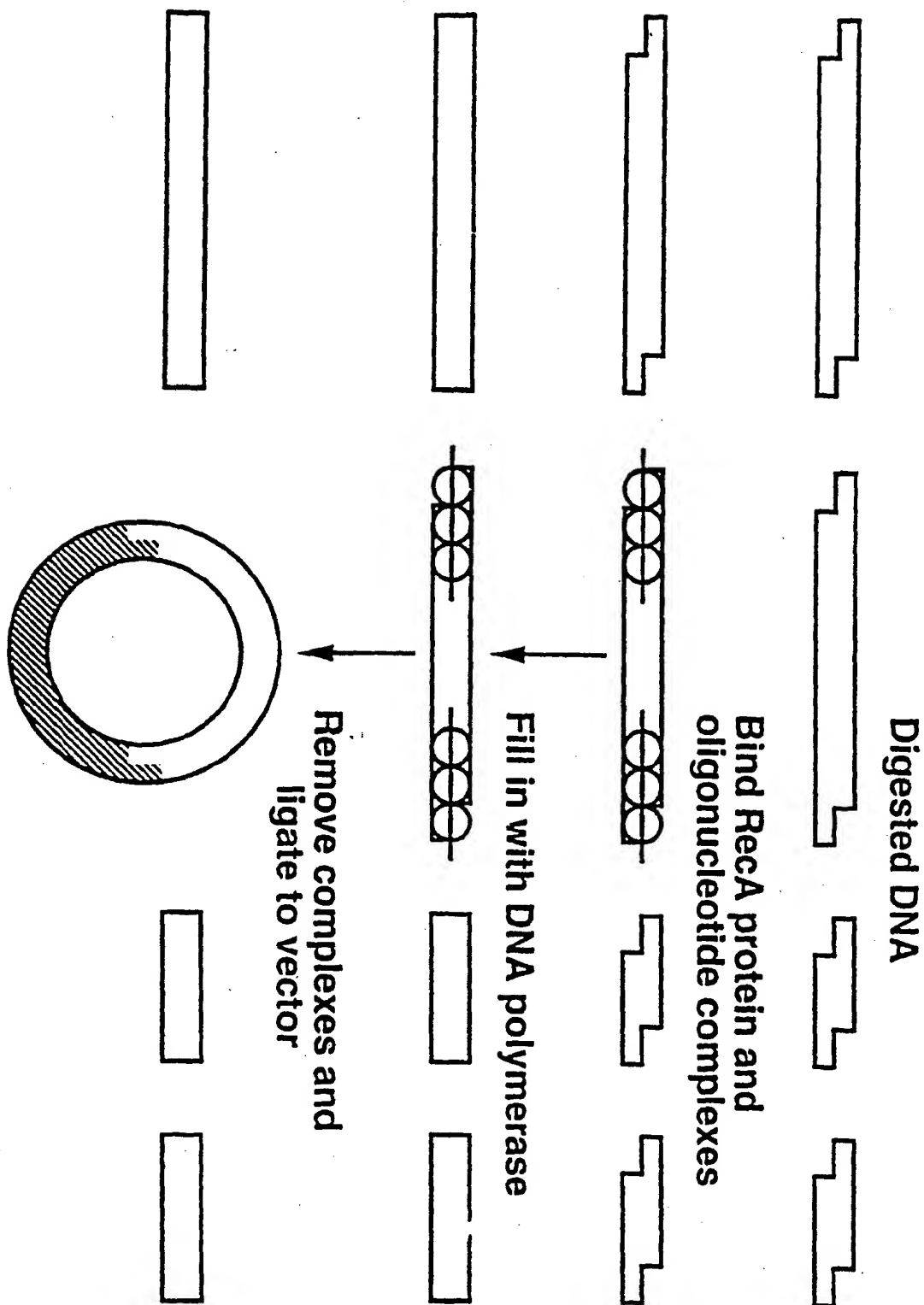
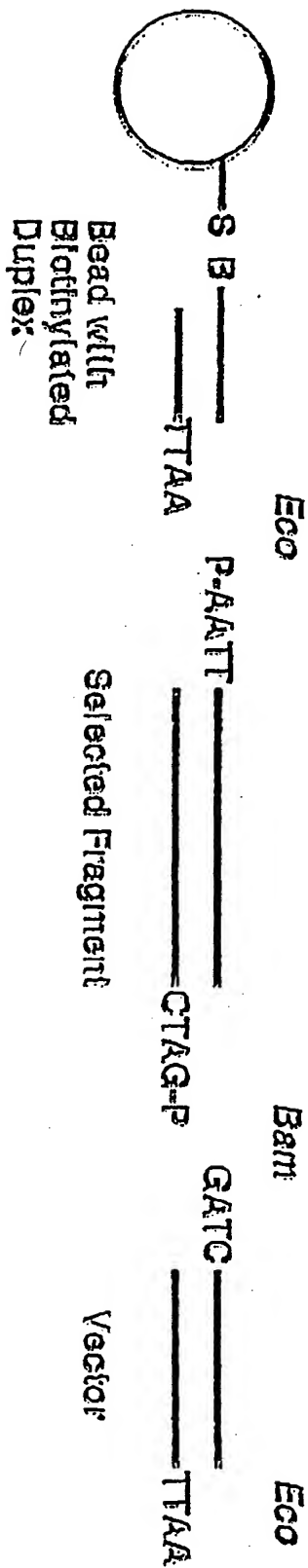


Figure 1

3 of 3

Figure 3



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 96/11721

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>SCIENCE, vol. 254, 1991, pages 1494-1497, XP002018783 L.J. FERRIN AND R.D. CAMERINI-OTERO: "Selective cleavage of human DNA: RecA-assisted restriction endonuclease (RARE) cleavage" *see the whole article*</p> <p>-----</p>	

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/66, C12Q 1/68</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 97/04111</b> <b>(43) International Publication Date:</b> 6 February 1997 (06.02.97)
<b>(21) International Application Number:</b> PCT/US96/11721 <b>(22) International Filing Date:</b> 16 July 1996 (16.07.96) <b>(30) Priority Data:</b> 60/001,384 21 July 1995 (21.07.95) US <b>(71) Applicant:</b> THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Office of Technology Transfer, Suite 325, 6011 Executive Boulevard, Rockville, MD 20852-3804 (US). <b>(72) Inventors:</b> FERRIN, Lance, Joseph; 8800 Alliston Hollow Way, Gaithersburg, MD 20879 (US). CAMERINI-OTERO, Rafael, Daniel; 3619 Littledale Road, Kensington, MD 20795 (US). <b>(74) Agent:</b> ALTMAN, Daniel, E.; 16th floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> RecA-ASSISTED CLONING OF DNA  <b>(57) Abstract</b>  DNA is cloned and labeled in a sequence-specific manner. The DNA is digested with one or more restriction enzymes which produce 3' recessed ends. A desired fragment is protected from elongation by DNA polymerase by addition of <i>E. coli</i> RecA protein and oligonucleotides about 30 to 60 bases in length complementary to the 3' recessed ends of the digested fragment. RecA and DNA polymerase are then inactivated, leaving only the desired fragment with 3' recessed ends which is then ligated into a vector containing complementary 3' recessed ends.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	México
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

## RecA-ASSISTED CLONING OF DNA

### Field of the Invention

The present invention relates to the sequence-specific cloning and labeling of DNA using the RecA protein. More specifically, the invention relates to the ability of RecA to selectively pair oligonucleotides to their homologous sequences at the 3' recessed ends of digested duplex DNA fragments and to protect these 3' ends from enzymatic conversion to blunt ends, thus facilitating cloning of a desired DNA fragment.

### Background of the Invention

The isolation and cloning of genomic DNA fragments is of paramount importance to the biomedical sciences. In this regard, several methods are available to amplify DNA and to isolate selected fragments in pure form. The most widely used amplification method is the polymerase chain reaction (PCR). In this method, oligonucleotide primers flanking a desired DNA sequence are used to amplify the sequence by repeated rounds of denaturation, annealing and extension steps. However, a major limitation of PCR is the small fragment size which may be reliably amplified, although recent improvements have allowed amplification of up to 22 kilobases (kb) (Cheng et al., *Proc. Natl. Acad. Sci. USA*, 91:5695, 1994; Foord et al., *PCR Methods and Applications*, 3:S149, 1994).

Other widely used methods of cloning genomic DNA fragments involve the construction and screening of DNA libraries, most commonly  $\lambda$  phage and cosmid vectors. Other vectors are now gaining widespread use for cloning large (> 100 kb) segments of DNA including yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs) and P1 phage derived artificial chromosomes (PACs). Such libraries, however, are difficult to construct and screen.

*E. coli* RecA protein has been used to screen libraries and to enrich for a selected DNA fragment (Rigas et al., *Proc. Natl. Acad. Sci. USA*, 83:9591, 1986; Honigberg et al., *Proc. Natl. Acad. Sci. USA*, 83:9586, 1986; Taidi-Laskowski et al., *Nucl. Acids. Res.*, 16:8157, 1988; Sena et al., *Nature Genet.*, 3:365, 1993; Jayasena et al., *J. Mol. Biol.*, 230:1015, 1993). These methods are based on the ability of RecA to specifically target single-stranded DNA to complementary target duplex DNA to create a three-stranded complex (Camerini-Otero et al., *Cell*, 73:217, 1993), or to pair two complementary single strands to the target duplex DNA to create a four-stranded complex. These strategies have not been applied to practical problems in molecular biology.

RecA-Assisted Restriction Endonuclease (RARE) Cleavage is a general and efficient method of targeting restriction enzyme cleavage to unique predetermined sites and is described in U.S. Patent Application Serial No. 08/088,910, the entire contents of which are hereby incorporated by reference, and by Ferrin et al. (*Nature Genet.*, 6:379, 1994). This method is based on the ability of RecA to pair oligonucleotides to homologous sequences in duplex DNA to form three-stranded complexes. These complexes protected the selected sites from enzymatic manipulation, and, after removal of the complexes, restriction enzyme cleavage was limited to the selected unmethylated sites. This method has been used to map and manipulate large segments of DNA (Ferrin, in *Genetic Engineering: Principles and Methods*, J. Setlow, Ed., Plenum Press, New York, 17:21-30, 1995; Barton et al., *Genes and Dev.*, 8:2453, 1994; Heineman et al., *J. Virol.*, 68:3317, 1994; Gourdon et al., *Nucl. Acids. Res.*, 22:4139, 1994).



Because of the practical size limitations of PCR cloning and the labor-intensive steps required in genomic DNA library construction and screening, there is a need for a simple, efficient method of labeling and cloning large fragments of genomic DNA. The present invention addresses this need.

#### Summary of the Invention

5 One aspect of the present invention is a method of cloning a genomic DNA fragment containing a predetermined DNA sequence. The method includes digesting DNA containing a predetermined DNA sequence with at least one restriction enzyme which generates 3' recessed ends to produce DNA fragments having 3' recessed ends. The DNA fragments are reacted with RecA protein and two oligonucleotides. These oligonucleotides are complementary to either DNA strand of the fragment containing the predetermined DNA sequence. In a preferred embodiment, the oligonucleotides  
10 are 30 to 60 nucleotides in length. The resulting fragments are then reacted with a DNA polymerase. As a result, all DNA fragments except the fragment containing the predetermined DNA sequence become blunt-ended. The oligonucleotides are dissociated from the ends of the fragment containing the predetermined DNA sequence. The DNA fragments are then ligated to a vector having 3' recessed ends complementary to those produced by the restriction enzyme. Only the fragment containing the predetermined DNA sequence is incorporated into the vector. The vector can  
15 be a plasmid, such as pBC SK<sup>+</sup> or pBS SK<sup>+</sup>. Advantageously, the vector is a yeast artificial chromosome, bacterial artificial chromosome or P1 phage artificial chromosome. Preferably, the restriction enzyme is *EcoRI* or a combination of *EcoRI* and *BamHI*. Advantageously, the DNA polymerase can be the exonuclease-deficient mutant of the Klenow fragment of *E. coli* DNA polymerase I.

The method can further comprise the step of size fractionating said DNA fragments of step (a) to enrich for  
20 the fragment containing the predetermined DNA sequence. This embodiment can further comprise, prior to the ligating step, ligating the enriched DNA fragments to a biotinylated duplex containing complementary 3' recessed ends, wherein the biotinylated duplex is bound to streptavidin-coated beads. In addition, the method can further comprise amplifying the DNA fragment containing the predetermined DNA sequence. Preferably, the amplifying step comprises transfection into bacteria or PCR.

25 The present invention also provides a method of diagnosing a genetic mutation in a mammal in which a variation of the above method is used. In this method, the fragment containing the fragment is amplified and it is determined if the mutation is present. Amplification can be by growth of the vector in a suitable microorganism or through PCR. Determination of the presence of the mutation can be accomplished by sequencing the fragment. Preferably, the mammal is a human and the DNA polymerase is the exonuclease-deficient mutant of the Klenow fragment of *E. coli* DNA  
30 polymerase I. The method can further comprise the step of size fractionating the DNA fragments of step (a) to enrich for the fragment containing the mutation. The method can also further comprise prior to the ligating step, ligating the enriched DNA fragments to a biotinylated duplex containing complementary 3' recessed ends, wherein the biotinylated duplex is bound to streptavidin-coated beads. In addition, the method can further comprise amplifying the DNA fragment containing the predetermined DNA sequence.

35 Another aspect of the present invention provides an article of manufacture which includes packaging material and at one or more reagents for cloning of DNA. The reagents for cloning of DNA includes *recA*, and the packaging

material includes instructions for using the reagents to clone DNA, such as by the method described above. The reagents can also include one or more restriction enzymes capable of generating 3' recessed ends, DNA polymerase and a vector having 3' cohesive ends. The 3' cohesive ends are preferably complementary to the 3' recessed ends generated by the restriction enzymes. In one embodiment, these restriction enzymes are *EcoRI* and *BamHI*. The DNA polymerase can be the Klenow fragment of *E. coli* DNA polymerase I, and the vector can be a plasmid.

#### Brief Description of the Drawings

Figure 1 is a schematic diagram of the strategy used for sequence-specific RecA-mediated amplification of DNA. The RecA-oligonucleotide complexes are indicated by the adjacent circles having a line passing through their centers.

Figure 2 is a schematic diagram of  $\lambda$  DNA digested with *HindIII*. The 2.3 kb fragment labeled by RecA-Assisted cloning is indicated by the arrow. The left (L) and right (R) 3' ends to which the L and R oligonucleotides are complementary are shown.

Figure 3 is a schematic diagram of the construct resulting from ligation of the 1.4 kb human *int-2* genomic DNA fragment to a vector and to a biotinylated DNA duplex bound to a streptavidin-coated bead. *EcoRI* and *BamHI* restriction sites and 4 base pair 5' overhangs are shown. S, streptavidin; B, biotin; and P, 5' phosphate.

#### Detailed Description of the Preferred Embodiments

The present invention relates to a method of sequence-specific genomic DNA cloning. This method is based on the ability of *E. coli* RecA protein to selectively pair oligonucleotides to their complementary sequences at the ends of duplex DNA. Genomic DNA is digested with one or two restriction enzymes which produce 3' recessed ends (5' overhangs). After addition of RecA protein and a pair of oligonucleotides, each complementary to one of the ends of a genomic DNA fragment of interest, the resulting three-stranded complexes become resistant to elongation by DNA polymerase and thus retain their 3' recessed ends after addition of the enzyme, while the unprotected genomic DNA fragments are filled in by the polymerase, thus becoming blunt-ended. Because most restriction endonucleases produce fragments having 3' recessed ends, these fragments were targeted for amplification using the method of the present invention. By using ligation conditions and vectors which greatly favor ligation to 3' recessed ends, protected fragments are selectively cloned. The vector into which the genomic DNA fragment is to be inserted is digested with the same restriction enzyme(s) as was the genomic DNA, or with a restriction enzyme(s) that produce the same 5' overhangs, resulting in complementary 3' recessed ends for insertion of the genomic DNA fragment.

The oligonucleotides used in the present method are complementary to a portion of the ends of a desired DNA fragment, including the 5' overhangs themselves. For example, digestion with *EcoRI* produces the 5' overhang TTAA. Thus, the oligonucleotides are complementary to this sequence plus additional sequence of the genomic DNA fragment adjacent to the overhang which is complementary to the remainder of the oligonucleotide. It will be appreciated that an oligonucleotide complementary to one DNA strand is identical in sequence to the other DNA strand. Of course, the oligonucleotides can be complementary to either strand at the end of the DNA duplex. Further, the oligonucleotides can also include additional nonhomologous sequences at their 5' or 3' ends; although these "nonhomologous tails" do not ordinarily increase the efficiency of protection. In addition, it is contemplated that only a single end of a desired DNA fragment can be protected with an oligonucleotide complementary thereto. In fact, a 200-fold enrichment of a particular

DNA sequence can be obtained by protection of one end of a DNA fragment. This entails cloning using a vector having one cohesive and one blunt end, or cutting the DNA fragment with another enzyme after the polymerase reaction.

The present invention has a number of important applications, including fast DNA cloning, DNA amplification in bacteria or by PCR, DNA sequence-based diagnostic tests and automated high-throughput DNA sequencing. The reagents for performing the method can also be supplied as a diagnostic kit for identification of mutations in a particular gene sequence. Genomic DNA can be digested with any single restriction enzyme which produces 3' recessed ends or any two different enzymes which produce 3' recessed ends. Such restriction enzymes are well known in the art and include, for example, *Bam*HI, *Eco*RI, *Hind*III, *Hinf*I, *Hpa*II, *Mlu*I and *Xba*I. In the preferred embodiment, oligonucleotides 30 to 60 bases in length are used, each having complete complementarity to the ends of the desired DNA fragment. Shorter and longer oligonucleotides can also be used, although reduced efficiency can sometimes result. A DNA polymerase and the four deoxynucleoside triphosphates are then added and allowed to fill in all available single-stranded sites with the exception of the fragment protected by the RecA-oligonucleotide complexes. In a preferred embodiment, The exonuclease-free Klenow Fragment (KF) of *E. coli* DNA polymerase I is used. This enzyme is efficient at creating blunt ends, can be added in excess without degrading DNA, is blocked by RecA-oligonucleotide complexes and is easily inactivated after completion of the reaction. The use of other DNA polymerases in the present method is also contemplated.

The RecA and KF enzymes are then inactivated, causing the RecA/oligonucleotide complex to dissociate from the DNA duplex. This inactivation can be accomplished by, for example, treatment with sodium dodecyl sulfate or phenol/chloroform extraction. The vast majority of the resulting DNA fragments in the mixture contain blunt ends as a result of the action of KF. However, the DNA fragment containing the predetermined DNA sequence of interest retains its 3' recessed ends as a result of the protection afforded by the RecA/oligonucleotide complex. These fragments can then be easily ligated into a vector having complementary 3' recessed ends.

After enrichment for a particular fragment, the fragment is ligated to a vector containing the appropriate 3' recessed ends. The insert is then amplified by, for example, transforming bacteria with the insert-containing plasmid or by PCR. Because DNA fragments having complementary 3' recessed ends are ligated to the vector much more readily than DNA fragments containing blunt ends, the resulting clones are highly enriched for the selected fragment.

In a preferred embodiment, in the cloning of genomic DNA fragments, it is desirable to size fractionate the digested DNA prior to the RecA/KF protection reaction to augment the enrichment of a particular fragment and to eliminate the cloning of small fragments. This can be accomplished by, for example, using an agarose gel followed by recovery of DNA from the relevant molecular weight region of the gel. In another preferred embodiment, the size fractionated DNA is ligated to both a short biotinylated duplex bound to streptavidin-coated beads which terminates with the same cohesive end as that produced by digestion with the restriction enzyme(s) used to digest the genomic DNA, and to a vector which has been digested with a restriction enzyme which produces 3' recessed ends complementary to those produced after digestion of the genomic DNA. Alternatively, the ligation reaction can be performed in two steps. In a preferred embodiment, the vector is a plasmid. Other vectors are also contemplated including bacteriophage vectors such as  $\lambda$ ; eukaryotic expression vectors such as the LacSwitch™ inducible mammalian expression system (Stratagene),

adenoviral vectors and the like. Particularly preferred vectors for the propagation of large DNA fragments include YACs, BACs and PACs. The vector is then used to transform cells which are expanded, resulting in amplification of the selected DNA fragment. Alternatively, the fragment can be amplified using PCR.

RecA-Assisted cloning has sufficient specificity to allow cloning directly from genomic DNA and is a much easier alternative than construction and screening of DNA libraries. The technique is preferable to PCR in the cloning of large (greater than about 5 kb) or highly repetitive fragments of DNA, especially if absolute fidelity is required due to the lower error rate of RecA-Assisted Cloning versus PCR.

RecA-Assisted Cloning can be used to identify specific mutations in a gene which give rise to genetic abnormalities and thus is useful in screening patients for genetic abnormalities or mutations which will predispose patients to certain conditions. Such mutations include point mutations, insertions and deletions. One particular use in this regard is in fetal screening. Fetal cells can be obtained by amniocentesis and analyzed for genetic defects including Tay-Sachs, sickle cell anemia,  $\beta$ -thalassemias, and any other desired genetic disease. Specific oligonucleotides are designed which will hybridize to the 3' ends of the fragment containing the DNA sequence of interest.

Many modifications of RecA-Assisted Cloning are contemplated. For example, the RecA/KF reactions worked well on DNA embedded in agarose which will be useful for molecules that would tend to shear in solution. For applications in which increased specificity is desired, RecA-Assisted Cloning can be used after RARE cleavage, or with type IIs restriction enzymes that create varied and asymmetric staggered ends unrelated to their recognition sites (Berger, *Anal. Biochem.*, 222:1, 1994). Increases in specificity would also be useful for labeling specific genomic DNA fragments using RecA-Assisted Cloning and is a viable alternative to detection methods such as Southern blotting. In addition, if conditions can be found that allow labeling of very short duplexes, the method can be a useful adjunct to sequencing by oligonucleotide array methods (Drmanac et al., *Science* 260:1649, 1993).

Sequence-specific labeling of a  $\lambda$  DNA fragment using RecA-Assisted Cloning was performed as described below.

#### Example 1

##### Sequence-specific labeling of a 2.3 kb $\lambda$ DNA fragment

*E. coli* RecA protein was prepared as described (Ferrin et al., *Science*, 254:1494, 1991) using an overproducing strain provided by Barbara McGrath of the Brookhaven National Laboratory, or purchased from Boehringer Mannheim (Indianapolis, IN). The sequence of the L oligonucleotide was 5'-gattatAGCTTTTCTAATTTAACCTTTGTCAGGTTACCA-3' (SEQ ID NO:1), and the R oligonucleotide was 5'-gattatAGCTTTGTGTGCCACCCACTACGACCTGCATAA-3' (SEQ ID NO:2). The lower case letters indicate sequences of nonhomologous tails, and the capital letters indicate the sequences of portions homologous to the ends of the  $\lambda$  fragment. Oligonucleotides over 30 bases in length were purified on acrylamide gels and concentrations were measured as described (Ferrin et al., *supra*).

The RecA protection reaction volume was 100  $\mu$ l and contained 25 mM Tris-acetate, pH 7.85, 4 mM magnesium acetate, 0.4 mM dithiothreitol, 0.5 mM spermidine, 1.1 mM ADP, 0.3 mM ATP- $\gamma$ -S (Fluka), 13  $\mu$ g of RecA protein, 0.32  $\mu$ g L or R oligonucleotide (or 0.16  $\mu$ g each of L and R), 2.5  $\mu$ g of *Hind*III-digested  $\lambda$  DNA (New England Biolabs, Beverly, MA) and 40  $\mu$ g bovine serum albumin (BSA; Sigma, St. Louis, MO), 38  $\mu$ M each of dATP, dCTP, dGTP and TTP, and 12.5 units of KF (United States Biochemical, Cleveland, OH). After a 10 minute incubation at 37°C, KF

and deoxynucleoside triphosphates were added and the reaction allowed to proceed for 30 minutes at 37°C. RecA and KF were then inactivated by extraction with phenol/chloroform (1:1), followed by extraction three times with diethyl ether, addition of sodium acetate to 0.3 M and precipitation with ethanol. The pellets were washed with 70% ethanol followed by ligation to the following short radioactive duplex for one hour at room temperature:

5                    5'-AGCTTACGATCGATGCCTTGACAT-3'                    (SEQ ID NO:3)  
                      3'-ATGCTAGCTACGGAAGTGTAGGAG-5'                    (SEQ ID NO:4)

The *HindIII* cohesive end is at the left, and the bottom strand was labeled with  $\gamma$ -<sup>32</sup>P-ATP using polynucleotide kinase. The kinase was heat-inactivated at 65°C for 10 minutes and the unreacted  $\gamma$ -<sup>32</sup>P-ATP was removed by gel  
 10 filtration (Chroma Spin + TE-10 columns; Clontech, Palo Alto, CA) prior to adding the top strand. The ligation reaction had a volume of 40  $\mu$ l and contained 1.0  $\mu$ g  $\lambda$  DNA, 0.8  $\mu$ g labeled duplex, 8 units of *E. coli* DNA ligase, and the buffer recommended by New England Biolabs without BSA. Excess duplex was removed by gel filtration followed by addition of bromphenol blue and glycerol. The samples were heated to 65°C for 3 minutes and analyzed by agarose gel  
 15 electrophoresis. Quantitation was performed using a Fuji Phosphor Imager and yields were calculated by comparison to the 2.3 kb band obtained from the reaction mixture containing the L oligonucleotide, but lacking KF after a small correction for a portion of the band removed by ligation to other fragments.

Efficient labeling of only the 2.3 kb band occurred when the *HindIII*-digested  $\lambda$  DNA was incubated with the L oligonucleotide, the R oligonucleotide and KF. In this case, the L and R oligonucleotides were used to protect both  
 20 the left and the right ends of the 2.3 kb  $\lambda$  DNA fragment followed by ligation of the short labeled duplex to both ends. When only the L or R oligonucleotide was used, each band on the agarose gel was only about half the intensity of the band obtained using both oligonucleotides. No specific labeling was observed if the ends were not protected (neither L nor R present), or when the restriction enzyme used to fragment the starting  $\lambda$  DNA produced blunt ends. In  
 addition, all of the fragments were labeled when KF was omitted.

The protection efficiency at each end of the 2.3 kb fragment was about 90%. Nonspecific protection of other  
 25 ends was detectable, but less than 0.5%, and labeling of the DNA with blunt ends was undetectable. Only 29 bases of sequence information at each end of the duplex was used in designing the oligonucleotides (33 bases if the 4 base single-stranded tail produced by *HindIII* is counted). A series of nine oligonucleotides was synthesized using an automated DNA synthesizer to investigate the parameters that determine protection efficiency. The efficiency was the same when the oligonucleotide contained 41 homologous bases, but dropped to 76% with 19 bases, and to less than  
 30 1% with 10 bases.

The oligonucleotides could have the same sequence at either strand at the end of the duplex without changing the efficiency. Addition of a tail that extended the oligonucleotide past the end of the fragment did not change the efficiency. These results were slightly more favorable than with RARE cleavage, and probably reflected the increased stability of complexes formed at the end of duplexes (Kim et al., *J. Mol. Biol.*, 247:874, 1995).

35                    To demonstrate RecA-assisted cloning using genomic DNA, we cloned a 1.4 kb *EcoRI-BamHI* fragment of the human *int-2* proto-oncogene as described in the following example.

## Example 2

RecA-assisted cloning of genomic DNA

The human *int-2* proto-oncogene has been mapped and sequenced (Casey et al., *Mol. Cell. Biol.*, 6:502, 1986; Brookes et al., *Oncogene*, 4:429, 1989). In this gene, one *EcoRI* site lies just upstream of exon 2 and, in about half of the alleles, a *BamHI* site is 6.9 kb downstream of the *EcoRI* site. Human genomic DNA was isolated from multiple placentas (Sigma), digested with *EcoRI* and *BamHI*, extracted with phenol/chloroform and ethanol precipitated as described in Example 1. Yields after the ethanol precipitation were typically about 60%. Digested DNA was size fractionated on a 0.8% SeaPlaque GTG (FMC BioProducts) agarose gel in TAE buffer (Sambrook et al., *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory Press, Plainview, NY, Second Edition, 1989). Multiple wells were loaded with 150-200  $\mu$ g DNA per well. The gels were run until the 1.4 kb fragment had migrated 4 to 6 cm. The marker lane was then removed, stained with ethidium bromide and used as a guide to excise 0.5 cm above and below the position of the 1.4 kb fragment. Size fractionation resulted in a modest sequence enrichment (about ten fold); a ten-fold decrease in the amount of reagents required; and eliminated small fragments that would preferentially be represented in the final clones.

DNA was extracted from the excised gel using GELase<sup>®</sup> (Epicentre) according to the manufacturer's directions; complete digestion was required for good yields. The yield of the complete protocol was 2 to 4%. Comparable yields were obtained with a silica gel extraction kit (Qiagen) or by electroelution (Pun et al., *Prep. Biochem.*, 20:123, 1990). The size of the extracted DNA from the heavily overloaded gels was checked on analytical gels. Depending on the amount available, DNA was quantified by absorbance, fluorescence (Labarca et al., *Anal. Biochem.*, 102:344, 1980), or spotting in an ethidium bromide solution (Sambrook et al., 1989).

The size-fractionated doubly digested human placental DNA was used as the starting DNA for the RecA protein/KF protection reaction. The conditions for this reaction were the same as described in Example 1 with the exception that the total volume was 1440  $\mu$ l and contained 3.2  $\mu$ g of each oligonucleotide, 360  $\mu$ g RecA protein, 2.8  $\mu$ g fractionated DNA, 570  $\mu$ g BSA and 450 units of KF. One nucleotide was identical to the *int-2* genomic sequence from 2290-2347: 5'-GGTCCGAGTGCGCGGAATTCGTCTCACTAAGACACTCCGGTTCTCTCAAAGCCAGGC-3'(SEQ ID NO:5), and the other was complementary to 3621-3677:

5'-TGGTCCTAGCTTGGATCCCATGTACCCTTGGCAAAGCATTCTACTGCCCACATCCCC-3'(SEQ ID NO:6). *EcoRI* and *BamHI* cleave 3' of residues 2304 and 3660, respectively (Casey et al., 1986; Brookes et al., 1989).

The protected fragments were ligated both to the pBS SK<sup>+</sup> vector (Stratagene, La Jolla, CA) and to DNA bound to streptavidin beads using T4 DNA ligase. This step reduces the number of clones containing only vector DNA. When plasmid or  $\lambda$  vectors were simply ligated to DNA from the RecA protein/KF reaction, the vast majority of clones did not contain an insert. Efforts were made to reduce this background by decreasing the vector concentration, but this also lowered the efficiency of the cloning procedure. Due to the low amount of the selected fragment in genomic DNA, a large concentration of vector facilitated the intermolecular vector-fragment ligation.

Ligation of the 1.4 kb fragment to the vector and to the biotinylated DNA duplex bound to the streptavidin-coated beads is schematically shown in Figure 3. Magnetic streptavidin beads (2 mg; Dynabeads M-280, Dynal) were

used according to the manufacturer's instructions and saturated with the following duplex that contained an *EcoRI* cohesive end that lacked a 5' phosphate:

5'-AATTCTACCAGAGGTACAAGGAGGA-3'

(SEQ ID NO:7)

3'-GATGGTCTCCATGTTCTCCTA-5'

(SEQ ID NO:8)

5 The oligonucleotide shown in SEQ ID NO:8 was synthesized with a biotin group at the 5' end using the LC Biotin-ON phosphormidite (Clontech). After binding, excess duplex was removed by washing the beads with 1 M NaCl, 50 mM Tris-HCl, pH 7.5, followed by T4 DNA ligase buffer (New England Biolabs). Vector was prepared by treating pBS SK<sup>+</sup> with *EcoRI*, *BamHI* and calf intestinal phosphatase. The small polylinker fragment arising from the digestion  
10 was removed by gel filtration. The ligation reaction contained the washed beads and 80  $\mu$ l of T4 DNA ligase buffer with 20% of the DNA from the RecA protein/KF protection reaction, 0.34  $\mu$ g of vector and 3,200 units of T4 DNA ligase. After 16 hours at 16°C, unligated DNA and vector were removed by washing the beads. The oligonucleotide shown in SEQ ID NO:7 (0.3  $\mu$ g) was added to replace any removed by washing. The *BamHI* site on the other side of the fragment was available to ligate to the vector.

15 Immobilized fragment-vector DNA was removed by treatment with 80 units of *EcoRI* in 100  $\mu$ l *EcoRI* buffer. The solution containing the fragment-vector DNA was removed from the beads, extracted and ethanol precipitated as described in Example 1, except that 20  $\mu$ g glycogen was added before the ethanol. To circularize the fragment-vector molecules, the DNA was treated with 1,600 units of T4 DNA ligase in 100  $\mu$ l of T4 ligase buffer for 16 hours at 16°C. The DNA was concentrated by ethanol precipitation and used to transform 50  $\mu$ l of *E. coli* XL-1 Blue MRF' (Stratagene)  
20 by electroporation in 0.1 cm cuvettes and a Gene Pulser apparatus (Bio-Rad, Richmond, CA). The cells were prepared for electroporation according to the instructions provided by Bio-Rad and yielded  $8 \times 10^5$  colonies per  $\mu$ g of plasmid DNA using standard 0.2 cm cuvettes. Cells were plated on Luria Broth (LB) agar containing ampicillin, tetracycline, isopropylthio- $\beta$ -D-galactoside (IPTG) and X-gal (Sambrook et al., 1989). Plasmid DNA was obtained by scraping and washing the resulting 7,000 white and 900 blue colonies from the plates with LB. Plasmid DNA was prepared using  
25 a kit (Qiagen) and extracted with cetyltrimethylammonium bromide (Ausubel, F., Ed., *Current Protocols in Molecular Biology*, Wiley, New York, 1995) to remove enzyme inhibitors.

Both the starting human placental DNA and the DNA obtained by RecA-Assisted Cloning were digested with *EcoRI* and *BamHI* and analyzed by agarose gel electrophoresis and Southern blotting. After staining with ethidium bromide, a large smear was observed in the lane containing digested human placental DNA, while only vector DNA was  
30 visible in the lane containing the DNA obtained by RecA-Assisted Cloning, although a faint smear of insert DNA centered at around 1.4 kb was observed on another gel containing twice as much DNA. The agarose gel was blotted onto charged nylon membranes using standard techniques, and the nylon membrane was probed with a labeled fragment (SS6) containing 0.6 kb of the selected *int-2* fragment (Casey et al., 1986; Brookes et al., 1989).

The amount of the *int-2* fragment was 20 times greater in the lane containing the RecA-cloned DNA as  
35 compared to the lane containing digested genomic DNA, even though the genomic DNA lane contained 80 times more DNA than the cloned DNA lane. Thus, a 1600-fold enrichment of the fragment was obtained by RecA-Assisted Cloning. These results reflected cloning of a fragment present at only about one copy per diploid human genome.

## Example 3

Multiple Cloning Trials

Multiple cloning trials were performed using both yeast and human DNA. With human DNA, the typical enrichment was 1,000 to 2,000-fold, and one *int-2* clone was present for every 2,000 to 4,000 colonies. At least one

5 *int-2* clone was obtained for every 70  $\mu$ g of starting genomic DNA. When the pooled DNA after one round of RecA-Assisted Cloning was subjected to the procedure a second time, 24% of the colonies contained the *int-2* fragment. The procedure was essentially identical to that described in Example 2, except that the RecA/KF reaction step was scaled down by a factor of 14. pBC SK<sup>+</sup> was used as the vector and clones were selected on chloramphenicol plates to eliminate any background from the previous plasmid vector which contained an ampicillin resistance gene. This

10 demonstrated an additional 500-fold enrichment and showed that incorrect clones arose mainly through a stochastic process, and not through a biased selection based on partial homology to the *int-2* sequence. A 1.2 kb *EcoRI-BamHI* yeast genomic DNA fragment containing the proximal portion of the RAD51 gene (Shinohara et al., *Cell* 69:457, 1992) was also cloned. The oligonucleotides used to clone this fragment had sequences complementary to positions 1-48:

5'-TGAAAATATTGAACAGTGAATAAAGCATAAAAAAAAAAATGTCGGATCCATAGCGCTAT-3 (SEQ ID NO:9), and 1164-1204:

15 5'-GGACTTACCTGTCCTGTCCTGAATTCACCGAAAAGCTCAGTAATAGAACCAGTTTCCACACC-3 (SEQ ID NO:10). The yeast genomic DNA was isolated as described by Ausubel (*ibid.* (p.13)). Conditions were identical to those described in Example 2, except that the RecA/KF reaction was scaled down by a factor of 14, pBC SK<sup>+</sup> was used as the vector and clones were selected on chloramphenicol plates.

Plasmid DNA from 10 *int-2* clones and 10 RAD51 clones were analyzed by restriction enzyme mapping. No

20 rearrangements were detected. The sequences of the two vector-insert junctions and about 400 bases of insert DNA of each of the 20 clones were determined. Plasmid DNA was prepared using a kit (Qiagen). Digestions were performed with *EcoRI* and *BamHI* and analyzed by electrophoresis on an agarose gel. Sequencing was performed on an Applied Biosystems model 373 sequencer using their PRISM DyeDeoxy Terminator Cycle Sequencing kit and the M13-20 and reverse primers. No clear deviations from the published sequences were detected, but as ambiguities in the sequences

25 occurred at a rate of about 1%, this could only be used to set an upper limit to the error rate of RecA-Assisted Cloning. One might expect the error rate to be closer to the *in vivo* error rate in *E. coli* of  $10^{-10}$  mutations/bp/chromosome duplication (Schaaper et al., *J. Biol. Chem.*, 268:23762, 1993; Drake, *Proc. Natl. Acad. Sci. USA*, 88:7160, 1991), rather than the PCR error rate of about  $10^{-5}$  (Barnes, *Proc. Natl. Acad. Sci. USA*, 91:2216, 1994).

Although the invention has been described with reference to particular preferred embodiments, the scope of the

30 invention is defined by the appended claims and should be construed to include reasonable equivalents.



## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: The Government of the United States, as represented by the Secretary, Department of Health and Human Services

(ii) TITLE OF INVENTION: RecA-Assisted Cloning of Genomic DNA

(iii) NUMBER OF SEQUENCES: 10

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Knobbe, Martens, Olson & Bear
- (B) STREET: 620 Newport Center Drive, Sixteenth Floor
- (C) CITY: Newport Beach
- (D) STATE: CA
- (E) COUNTRY: USA
- (F) ZIP: 92660

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Altman, Daniel E.
- (B) REGISTRATION NUMBER: 34,115
- (C) REFERENCE/DOCKET NUMBER: NIH117.001PR

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (714) 760-0404
- (B) TELEFAX: (714) 760-9502

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GATTATAGCT TTTCTAATTT AACCTTTGTC AGGTTACCA

39

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GATTATAGCT TTGTGTGCCA CCCACTACGA CCTGCATAA

39

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCCTACGAT CGATGCCTTG ACAT

24

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATGCTAGCTA CGGAAGTGTG GGAG

24

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 58 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGTCCGAGTG CGCGGAATTC GTCTCACTAA GACACTCCGG TTCTCTCCAA AGCCAGGC

58

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 57 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TGGTCCTAGC TTGGATCCCA TGTACCCTTG GCAAAGCATT CTACTGCCCA CATCCEC

57

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AATTCTACCA GAGGTACAAG GAGGA

25

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GATGGTCTCC ATGTTCTCTCC TA

22

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGAAAATATT GAACAGTGAA TAAAGCATAA AAAAAAATG TCGGATCCAT AGCGCTAT

58

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 57 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGACTTACCT GTCCTGAATT CACCGAAAAG CTCAGTAATA GAACCAGTTT CCACACC

57

WHAT IS CLAIMED IS:

1. A method of cloning a DNA fragment containing a predetermined DNA sequence, comprising the steps of:
  - (a) digesting DNA containing a predetermined DNA sequence with at least one restriction enzyme which generates 3' recessed ends to produce DNA fragments having 3' recessed ends;
  - (b) reacting said DNA fragments with RecA protein and two oligonucleotides, said oligonucleotides being complementary to either DNA strand of the fragment containing the predetermined DNA sequence;
  - (c) reacting the DNA fragments resulting from step (b) with a DNA polymerase, whereby all DNA fragments except the fragment containing the predetermined DNA sequence become blunt-ended;
  - (d) dissociating said oligonucleotides from the ends of the fragment containing the predetermined DNA sequence; and
  - (e) ligating said DNA fragments to a vector having 3' recessed ends complementary to those produced by the restriction enzyme, whereby only the fragment containing the predetermined DNA sequence is incorporated into said vector.
2. The method of Claim 1, wherein said oligonucleotides are between 30 and 60 bases in length.
3. The method of Claim 1, wherein said restriction enzyme is *EcoRI*.
4. The method of Claim 1, wherein two restriction enzymes are reacted in step (a).
5. The method of Claim 4, wherein said restriction enzymes are *EcoRI* and *BamHI*.
6. The method of Claim 1, wherein said DNA polymerase is the Klenow fragment of *E. coli* DNA polymerase I.
7. The method of Claim 1, wherein said vector is a plasmid.
8. The method of Claim 7, wherein said plasmid is pBC SK<sup>+</sup> or pBS SK<sup>+</sup>.
9. The method of Claim 1, wherein said vector is a yeast artificial chromosome, bacterial artificial chromosome or P1 phage derived artificial chromosome.
10. The method of Claim 1, further comprising the step of size fractionating said DNA fragments of step (a) to enrich for the fragment containing the predetermined DNA sequence.
11. The method of Claim 10, further comprising, prior to the ligating step, ligating the enriched DNA fragments to a biotinylated duplex containing complementary 3' recessed ends, wherein said biotinylated duplex is bound to streptavidin-coated beads.
12. The method of Claim 1, further comprising amplifying the DNA fragment containing the predetermined DNA sequence.
13. The method of Claim 12, wherein said amplifying step comprises transfection into bacteria.
14. The method of Claim 12, wherein said amplifying step comprises PCR.
15. A method of diagnosing a genetic mutation in a mammal, comprising the steps of:
  - (a) isolating genomic DNA containing said mutation from a mammal;

(b) digesting said genomic DNA with one or more restriction enzymes which generate 3' recessed ends to produce genomic DNA fragments having 3' recessed ends;

(c) reacting said genomic DNA fragments with RecA protein and two oligonucleotides, said oligonucleotides being complementary to the ends of the fragment containing the mutation;

5 (c) reacting the genomic DNA fragments resulting from step (b) with a DNA polymerase, whereby all genomic DNA fragments except the fragment containing the mutation become blunt-ended;

(d) dissociating said oligonucleotides from the ends of the fragment containing the mutation;

(e) ligating said DNA fragments to a vector having 3' recessed ends complementary to those produced by the restriction enzyme(s), whereby only the fragment containing the mutation is incorporated into said vector;

10 (f) amplifying the fragment containing the mutation; and

(g) determining whether the mutation is present in the amplified fragments.

16. The method of Claim 15, wherein step (f) comprises growth of said vector in a suitable microorganism.

17. The method of Claim 15, wherein step (f) comprises PCR.

18. The method of Claim 15, wherein said oligonucleotides are between 30 and 60 bases in length.

15 19. The method of Claim 15, wherein step (g) comprises sequencing said fragment.

20. The method of Claim 15, wherein step (d) comprises treating with sodium dodecyl sulfate or phenol/chloroform.

21. The method of Claim 15, wherein said mammal is a human.

22. The method of Claim 15, wherein said DNA polymerase is the exonuclease-deficient Klenow fragment  
20 of *E. coli* DNA polymerase I.

23. The method of Claim 15, wherein said vector is a plasmid.

24. The method of Claim 15, wherein said vector is a yeast artificial chromosome, bacterial artificial chromosome or P1 phage derived artificial chromosome.

25. The method of Claim 15, further comprising the step of size fractionating said DNA fragments of step  
25 (a) to enrich for the fragment containing the mutation.

26. The method of Claim 25, further comprising prior to the ligating step, ligating the enriched DNA fragments to a biotinylated duplex containing complementary cohesive ends, wherein said biotinylated duplex is bound to streptavidin-coated beads.

27. An article of manufacture comprising packaging material and at one or more reagents for cloning of  
30 DNA, wherein the reagents for cloning of DNA comprise recA, and wherein the packaging material comprises instructions for using the reagents to clone DNA.

28. An article of manufacture comprising packaging material and at one or more reagents for cloning of DNA, wherein the reagents for cloning of DNA comprise recA protein, and wherein the packaging material comprises instructions for using the reagents to clone DNA according to the method of Claim 1.

35 29. An article of manufacture according to Claim 27, wherein the instructions for using the reagents comprise instructions to identify one or more mutations in genomic DNA.

30. An article of manufacture according to Claim 27, wherein the reagents additionally comprise one or more restriction enzymes capable of generating 3' recessed ends.
31. An article of manufacture according to Claim 27, wherein the reagents additionally comprise DNA polymerase.
- 5 32. An article of manufacture according to Claim 27, wherein the reagents additionally comprise a vector having 3' cohesive ends.
33. An article of manufacture according to Claim 30, wherein the reagents additionally comprise a vector having 3' cohesive ends, and wherein the 3' cohesive ends are complementary to the 3' recessed ends generated by the restriction enzymes.
- 10 34. An article of manufacture according to Claim 30, wherein the restriction enzymes comprise *EcoRI*.
35. An article of manufacture according to Claim 34, wherein the restriction enzymes comprise *EcoRI* and *BamHI*.
36. An article of manufacture according to Claim 31, wherein the DNA polymerase is the Klenow fragment  
15 of *E. coli* DNA polymerase I.
37. An article of manufacture according to Claim 32, wherein said vector is a plasmid.

1 of 3

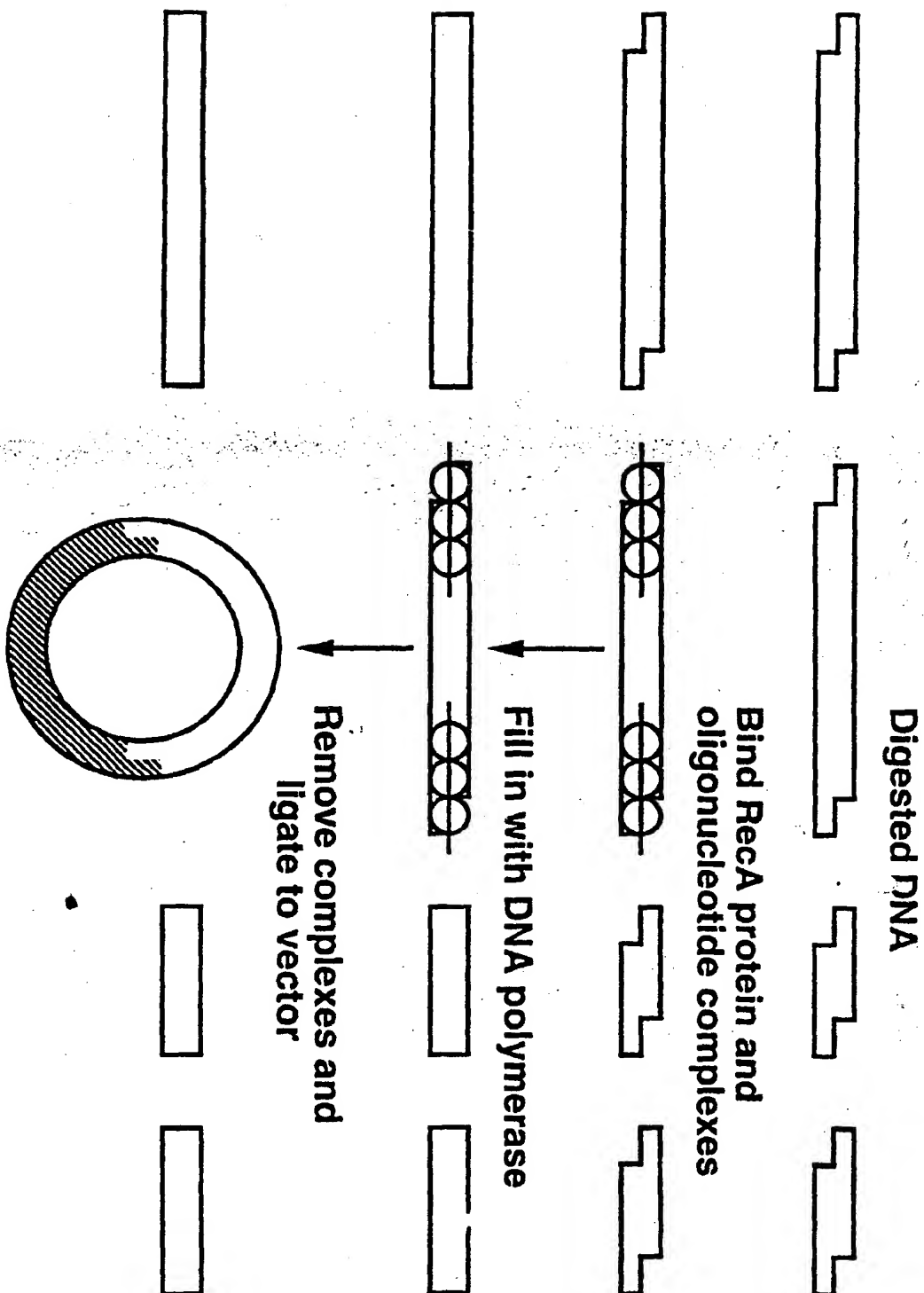
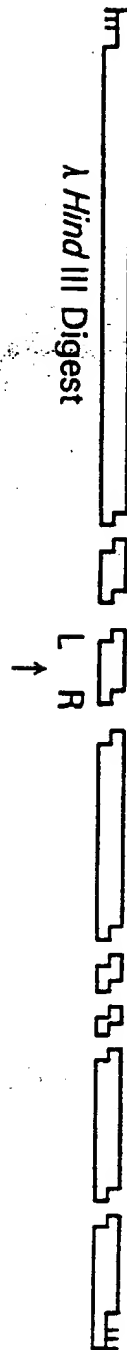


Figure 1



2 of 3

Figure 2



3 of 3

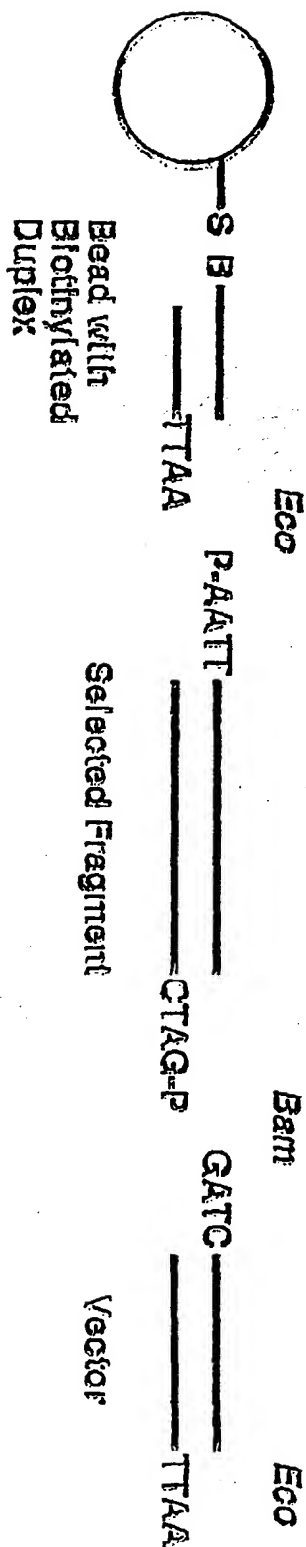


Figure 3

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 96/11721

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 C12N15/66 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BIOTECHNIQUES, vol. 10, no. 1, 1991, pages 62-68, XP002018782 D.H. JONES AND B.H. HOWARD: "A rapid method for recombination and site-specific mutagenesis by placing homologous ends on DNA using polymerase chain reaction" *see the whole article* ---	
A	WO,A,92 08791 (THE UNITED STATES OF AMERICA) 29 May 1992 *see the whole patent* ---	

-/--

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search

18 November 1996

Date of mailing of the international search report

06.12.96

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Marie, A

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 96/11721

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>SCIENCE, vol. 254, 1991, pages 1494-1497, XP002018783 L.J. FERRIN AND R.D. CAMERINI-OTERO: "Selective cleavage of human DNA: RecA-assisted restriction endonuclease (RARE) cleavage" *see the whole article*</p> <p style="text-align: center;">-----</p>	

# INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/US 96/11721

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9208791	29-05-92	AU-B- 658386	13-04-95
		AU-A- 9111691	11-06-92
		CA-A- 2095624	10-05-92
		EP-A- 0556330	25-08-93
		JP-T- 6500926	27-01-94
		US-A- 5460941	24-10-95
		US-A- 5510473	23-04-96
-----			